Are Declustered Earthquake Catalogs Poisson?

Philip B. Stark
Department of Statistics, UC Berkeley
Brad Luen
Department of Mathematics, Reed College

14 October 2010
Department of Statistics, Penn State University

Outline

Abstract

Claims that the times of events in declustered catalogs of Southern California seismicity fit Poisson process models have used tests that ignore earthquake locations. They divide time into intervals, count the events in the intervals, and then apply a chi-square test to those counts, calculating the expected number of intervals with each count from a Poisson distribution. The chi-square statistic does is insensitive to the order of the counts. Other tests give strong evidence that declustered Southern California catalogs do not follow a homogeneous temporal Poisson process. Spatial information is also telling: For declustered Southern California Earthquake Center catalogs for 1932–1971 and 2009, an abstract permutation test gives evidence that event times and locations are not conditionally exchangeable, a necessary condition for them to follow a spatially heterogeneous, temporally homogeneous Poisson process.

Phenomenology

- Earthquakes destroy and kill. Studied since ancient times. Prediction is an old goal: save lives and property.

- Phenomenology good. Physics not well understood.

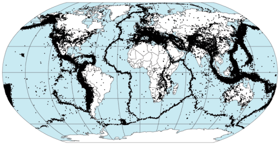
Phenomenology

- Earthquakes destroy and kill. Studied since ancient times. Prediction is an old goal: save lives and property.

- Phenomenology good. Physics not well understood.

Quake Physics versus Quake Statistics

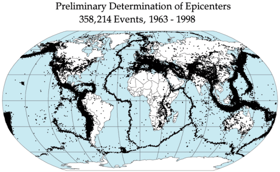
- Distribution in space, clustering in time, distribution of sizes (Gutenberg-Richter law: $N \propto 10^{a-bM}$)



- Foreshocks, aftershocks, swarms—no physics-based definitions
- Clustering makes *some* prediction easy: If there's a big quake, predict that there will be another, close and soon. Not very useful.
- Physics hard: Quakes are gnat's whiskers on Earth's tectonic energy budget
- Spatiotemporal Poisson model doesn't fit
- More complex models "motivated by physics"

Quake Physics versus Quake Statistics

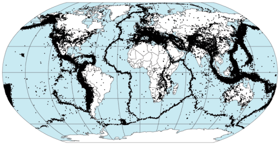
- Distribution in space, clustering in time, distribution of sizes (Gutenberg-Richter law: $N \propto 10^{a-bM}$)



- Foreshocks, aftershocks, swarms—no physics-based definitions
- Clustering makes *some* prediction easy: If there's a big quake, predict that there will be another, close and soon. Not very useful.
- Physics hard: Quakes are gnat's whiskers on Earth's tectonic energy budget
- Spatiotemporal Poisson model doesn't fit
- More complex models "motivated by physics"

Quake Physics versus Quake Statistics

- Distribution in space, clustering in time, distribution of sizes (Gutenberg-Richter law: $N \propto 10^{a-bM}$)

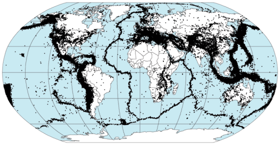


- Foreshocks, aftershocks, swarms—no physics-based definitions
- Clustering makes *some* prediction easy: If there's a big quake, predict that there will be another, close and soon. Not very useful.
- Physics hard: Quakes are gnat's whiskers on Earth's tectonic energy budget
- Spatiotemporal Poisson model doesn't fit
- More complex models "motivated by physics"

Quake Physics versus Quake Statistics

- Distribution in space, clustering in time, distribution of sizes (Gutenberg-Richter law: $N \propto 10^{a-bM}$)

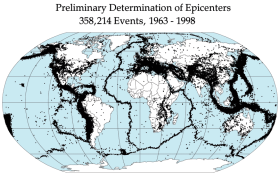
Preliminary Determination of Epicenters
358,214 Events, 1963 - 1998



- Foreshocks, aftershocks, swarms—no physics-based definitions
- Clustering makes *some* prediction easy: If there's a big quake, predict that there will be another, close and soon. Not very useful.
- Physics hard: Quakes are gnat's whiskers on Earth's tectonic energy budget
- Spatiotemporal Poisson model doesn't fit
- More complex models "motivated by physics"

Quake Physics versus Quake Statistics

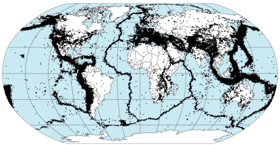
- Distribution in space, clustering in time, distribution of sizes (Gutenberg-Richter law: $N \propto 10^{a-bM}$)



- Foreshocks, aftershocks, swarms—no physics-based definitions
- Clustering makes *some* prediction easy: If there's a big quake, predict that there will be another, close and soon. Not very useful.
- Physics hard: Quakes are gnat's whiskers on Earth's tectonic energy budget
- Spatiotemporal Poisson model doesn't fit
- More complex models "motivated by physics"

Quake Physics versus Quake Statistics

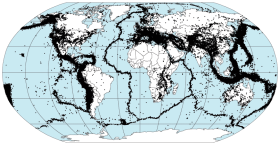
- Distribution in space, clustering in time, distribution of sizes (Gutenberg-Richter law: $N \propto 10^{a-bM}$)



- Foreshocks, aftershocks, swarms—no physics-based definitions
- Clustering makes *some* prediction easy: If there's a big quake, predict that there will be another, close and soon. Not very useful.
- Physics hard: Quakes are gnat's whiskers on Earth's tectonic energy budget
- Spatiotemporal Poisson model doesn't fit
- More complex models "motivated by physics"

Quake Physics versus Quake Statistics

- Distribution in space, clustering in time, distribution of sizes (Gutenberg-Richter law: $N \propto 10^{a-bM}$)



- Foreshocks, aftershocks, swarms—no physics-based definitions
- Clustering makes *some* prediction easy: If there's a big quake, predict that there will be another, close and soon. Not very useful.
- Physics hard: Quakes are gnat's whiskers on Earth's tectonic energy budget
- Spatiotemporal Poisson model doesn't fit
- More complex models "motivated by physics"

Why decluster?

Online FAQ for USGS Earthquake Probability Mapping Application:

Q: "Ok, so why do you decluster the catalog?"

A: "to get the best possible estimate for the rate of mainshocks"

"the methodology requires a catalog of independent events (Poisson model), and declustering helps to achieve independence."

- What's a mainshock?
- Aren't foreshocks and aftershocks potentially destructive?

Why decluster?

Online FAQ for USGS Earthquake Probability Mapping Application:

Q: "Ok, so why do you decluster the catalog?"

A: "to get the best possible estimate for the rate of mainshocks"

"the methodology requires a catalog of independent events (Poisson model), and declustering helps to achieve independence."

- What's a mainshock?
- Aren't foreshocks and aftershocks potentially destructive?

Why decluster?

Online FAQ for USGS Earthquake Probability Mapping Application:

Q: “Ok, so why do you decluster the catalog?”

A: “to get the best possible estimate for the rate of mainshocks”

“the methodology requires a catalog of independent events (Poisson model), and declustering helps to achieve independence.”

- What’s a mainshock?
- Aren’t foreshocks and aftershocks potentially destructive?

Why decluster?

Online FAQ for USGS Earthquake Probability Mapping Application:

Q: “Ok, so why do you decluster the catalog?”

A: “to get the best possible estimate for the rate of mainshocks”

“the methodology requires a catalog of independent events (Poisson model), and declustering helps to achieve independence.”

- What’s a mainshock?
- Aren’t foreshocks and aftershocks potentially destructive?

“Main events,” “foreshocks,” and “aftershocks”

- An event that the declustering method does not remove is a main shock.
- An event that the declustering method removes is a foreshock or an aftershock.

  . . . profound shrug . . .

  Where’s the physics?

Declustering Methods

- Window-based methods
  - Main-shock window: punch hole in catalog near each “main shock”
  - Linked window: every event has a window.
    Clusters are maximal sets of events such that each is in the window of some other event in the group.
    Replace cluster by single event: first, largest, “equivalent”

  Generally, larger events have larger space-time windows
- Stochastic methods: use chance to decide which events to keep
- Other methods (e.g., waveform similarity)

Declustering Methods

- Window-based methods
  - Main-shock window: punch hole in catalog near each “main shock”
  - Linked window: every event has a window.
    Clusters are maximal sets of events such that each is in the window of some other event in the group.
    Replace cluster by single event: first, largest, “equivalent”

  Generally, larger events have larger space-time windows
- Stochastic methods: use chance to decide which events to keep
- Other methods (e.g., waveform similarity)

Declustering Methods

- Window-based methods
  - Main-shock window: punch hole in catalog near each “main shock”
  - Linked window: every event has a window.
    Clusters are maximal sets of events such that each is in the window of some other event in the group.
    Replace cluster by single event: first, largest, “equivalent”

  Generally, larger events have larger space-time windows
- Stochastic methods: use chance to decide which events to keep
- Other methods (e.g., waveform similarity)

Are “main events” Poisson in time?

Gardner & Knopoff, 1974:
“Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian?”

Abstract: “Yes.”

Aftershocks: defined as above

Statistical test: chi-square using counts of events in intervals

Easy to make declustered catalogs indistinguishable from Poisson by deleting enough shocks—or by using a weak test. Shrug.

Are “main events” Poisson in time?

Gardner & Knopoff, 1974:
“Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian?”

Abstract: “Yes.”

Aftershocks: defined as above

Statistical test: chi-square using counts of events in intervals

Easy to make declustered catalogs indistinguishable from Poisson by deleting enough shocks—or by using a weak test. Shrug.

Are “main events” Poisson in time?

Gardner & Knopoff, 1974:
“Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian?”

Abstract: “Yes.”

Aftershocks: defined as above

Statistical test: chi-square using counts of events in intervals

Easy to make declustered catalogs indistinguishable from Poisson by deleting enough shocks—or by using a weak test.
Shrug.

Chi-square test

- Pick K. Partition the study period into K time intervals.
- n: total number of events. N_k: events in interval k.
- Pick $B \geq 2$, the number of "bins."
  For $b \in \{0, \ldots, B-2\}$, O_b is the number of intervals that contain b events.
  O_{B-1} is the number of intervals with $B-1$ or more events.
- Estimate the rate of events by $\hat{\lambda} = n/K$.
- Set $E_b \equiv K e^{-\hat{\lambda}} \frac{\hat{\lambda}^b}{b!}$ for $b = 0, \ldots, B-2$, and set $E_{B-1} \equiv K - \sum_{b=0}^{B-2} E_b$.
- Chi-square statistic: $\chi^2 \equiv \sum_{b=0}^{B-1} \frac{(O_b - E_b)^2}{E_b}$.
  Nominal P-value: tail area of chi-square distribution, d.f. $= d$.
  Depends on K, B, d, and method of estimating λ.

Chi-square test

- Pick K. Partition the study period into K time intervals.
- n: total number of events. N_k: events in interval k.
- Pick $B \geq 2$, the number of "bins."
  For $b \in \{0, \ldots, B-2\}$, O_b is the number of intervals that contain b events.
  O_{B-1} is the number of intervals with $B-1$ or more events.
- Estimate the rate of events by $\hat{\lambda} = n/K$.
- Set $E_b \equiv K e^{-\hat{\lambda}} \frac{\hat{\lambda}^b}{b!}$ for $b = 0, \ldots, B-2$, and set $E_{B-1} \equiv K - \sum_{b=0}^{B-2} E_b$.
- Chi-square statistic: $\chi^2 \equiv \sum_{b=0}^{B-1} \frac{(O_b - E_b)^2}{E_b}$.
  Nominal P-value: tail area of chi-square distribution, d.f. $= d$.
  Depends on K, B, d, and method of estimating λ.

Chi-square test

- Pick K. Partition the study period into K time intervals.
- n: total number of events. N_k: events in interval k.
- Pick $B \geq 2$, the number of "bins."
  For $b \in \{0, \ldots, B-2\}$, O_b is the number of intervals that contain b events.
  O_{B-1} is the number of intervals with $B-1$ or more events.
- Estimate the rate of events by $\hat{\lambda} = n/K$.
- Set $E_b \equiv K e^{-\hat{\lambda}} \frac{\hat{\lambda}^b}{b!}$ for $b = 0, \ldots, B-2$, and set $E_{B-1} \equiv K - \sum_{b=0}^{B-2} E_b$.
- Chi-square statistic: $\chi^2 \equiv \sum_{b=0}^{B-1} \frac{(O_b - E_b)^2}{E_b}$.
  Nominal P-value: tail area of chi-square distribution, d.f. $= d$.
  Depends on K, B, d, and method of estimating λ.

Chi-square test

- Pick K. Partition the study period into K time intervals.
- n: total number of events. N_k: events in interval k.
- Pick $B \geq 2$, the number of "bins."
  For $b \in \{0, \ldots, B-2\}$, O_b is the number of intervals that contain b events.
  O_{B-1} is the number of intervals with $B-1$ or more events.
- Estimate the rate of events by $\hat{\lambda} = n/K$.
- Set $E_b \equiv K e^{-\hat{\lambda}} \frac{\hat{\lambda}^b}{b!}$ for $b = 0, \ldots, B-2$, and set $E_{B-1} \equiv K - \sum_{b=0}^{B-2} E_b$.
- Chi-square statistic: $\chi^2 \equiv \sum_{b=0}^{B-1} \frac{(O_b - E_b)^2}{E_b}$.
  Nominal P-value: tail area of chi-square distribution, d.f. $= d$.
  Depends on K, B, d, and method of estimating λ.

Chi-square test

- Pick K. Partition the study period into K time intervals.
- n: total number of events. N_k: events in interval k.
- Pick $B \geq 2$, the number of "bins."
  For $b \in \{0, \ldots, B-2\}$, O_b is the number of intervals that contain b events.
  O_{B-1} is the number of intervals with $B-1$ or more events.
- Estimate the rate of events by $\hat{\lambda} = n/K$.
- Set $E_b \equiv K e^{-\hat{\lambda}} \frac{\hat{\lambda}^b}{b!}$ for $b = 0, \ldots, B-2$, and set $E_{B-1} \equiv K - \sum_{b=0}^{B-2} E_b$.
- Chi-square statistic: $\chi^2 \equiv \sum_{b=0}^{B-1} \frac{(O_b - E_b)^2}{E_b}$.
  Nominal P-value: tail area of chi-square distribution, d.f. $= d$.
  Depends on K, B, d, and method of estimating λ.

Chi-square test

- Pick K. Partition the study period into K time intervals.
- n: total number of events. N_k: events in interval k.
- Pick $B \geq 2$, the number of "bins."
  For $b \in \{0, \ldots, B-2\}$, O_b is the number of intervals that contain b events.
  O_{B-1} is the number of intervals with $B-1$ or more events.
- Estimate the rate of events by $\hat{\lambda} = n/K$.
- Set $E_b \equiv K e^{-\hat{\lambda}} \frac{\hat{\lambda}^b}{b!}$ for $b = 0, \ldots, B-2$, and set $E_{B-1} \equiv K - \sum_{b=0}^{B-2} E_b$.
- Chi-square statistic: $\chi^2 \equiv \sum_{b=0}^{B-1} \frac{(O_b - E_b)^2}{E_b}$.
  Nominal P-value: tail area of chi-square distribution, d.f. $= d$.
  Depends on K, B, d, and method of estimating λ.

Gardner-Knopoff, 1974

Chi-square test on a number of declustered catalogs, including a catalog of earthquakes with $M \geq 3.8$ in the Southern California, 1932–1971.

Raw: 1,751 events.
Close to SCEC catalog for 1932–1971, but not an exact match (1,556 events w/ $M \geq 3.8$; see below)

Declustered catalog: 503 events.

10-day intervals.

$d = 2$ degrees of freedom.

Don't give B; don't explain how they estimated λ.

Chi-square approximation

Null for simple chi-square test: data are multinomial with known category probabilities. Here, requires

(i) $\Pr\{N_k = b\}_{b=1}^{B}$ known and don't depend on k

(ii) $\{N_k\}_{k=1}^{K}$ are iid.

Neither is true.

- Bin probabilities are estimated. (Asymptotic justification for MLE based on category frequencies; apparently not what is done.)
- "Trial" corresponds to an interval. Category based the number of shocks in the interval.
- Condition on total number of shocks to estimate the rate: dependence among trials. Joint distribution not multinomial.

Chi-square approximation

Null for simple chi-square test: data are multinomial with known category probabilities. Here, requires

(i) $\Pr\{N_k = b\}_{b=1}^{B}$ known and don't depend on k

(ii) $\{N_k\}_{k=1}^{K}$ are iid.

Neither is true.

- Bin probabilities are estimated. (Asymptotic justification for MLE based on category frequencies; apparently not what is done.)
- "Trial" corresponds to an interval. Category based the number of shocks in the interval.
- Condition on total number of shocks to estimate the rate: dependence among trials. Joint distribution not multinomial.

Chi-square approximation

Null for simple chi-square test: data are multinomial with known category probabilities. Here, requires

(i) $\Pr\{N_k = b\}_{b=1}^{B}$ known and don't depend on k

(ii) $\{N_k\}_{k=1}^{K}$ are iid.

Neither is true.

- Bin probabilities are estimated. (Asymptotic justification for MLE based on category frequencies; apparently not what is done.)
- "Trial" corresponds to an interval. Category based the number of shocks in the interval.
- Condition on total number of shocks to estimate the rate: dependence among trials. Joint distribution not multinomial.

Chi-square approximation

Null for simple chi-square test: data are multinomial with known category probabilities. Here, requires

(i) $\Pr\{N_k = b\}_{b=1}^{B}$ known and don't depend on k

(ii) $\{N_k\}_{k=1}^{K}$ are iid.

Neither is true.

- Bin probabilities are estimated. (Asymptotic justification for MLE based on category frequencies; apparently not what is done.)
- "Trial" corresponds to an interval. Category based the number of shocks in the interval.
- Condition on total number of shocks to estimate the rate: dependence among trials. Joint distribution not multinomial.

Chi-square approximation

Null for simple chi-square test: data are multinomial with known category probabilities. Here, requires

(i) $\Pr\{N_k = b\}_{b=1}^{B}$ known and don't depend on k

(ii) $\{N_k\}_{k=1}^{K}$ are iid.

Neither is true.

- Bin probabilities are estimated. (Asymptotic justification for MLE based on category frequencies; apparently not what is done.)
- "Trial" corresponds to an interval. Category based the number of shocks in the interval.
- Condition on total number of shocks to estimate the rate: dependence among trials. Joint distribution not multinomial.

Chi-square approximation

Null for simple chi-square test: data are multinomial with known category probabilities. Here, requires

(i) $\Pr\{N_k = b\}_{b=1}^{B}$ known and don't depend on k

(ii) $\{N_k\}_{k=1}^{K}$ are iid.

Neither is true.

- Bin probabilities are estimated. (Asymptotic justification for MLE based on category frequencies; apparently not what is done.)
- "Trial" corresponds to an interval. Category based the number of shocks in the interval.
- Condition on total number of shocks to estimate the rate: dependence among trials. Joint distribution not multinomial.

Chi-square approximation

Null for simple chi-square test: data are multinomial with known category probabilities. Here, requires

(i) $\Pr\{N_k = b\}_{b=1}^{B}$ known and don't depend on k

(ii) $\{N_k\}_{k=1}^{K}$ are iid.

Neither is true.

- Bin probabilities are estimated. (Asymptotic justification for MLE based on category frequencies; apparently not what is done.)
- "Trial" corresponds to an interval. Category based the number of shocks in the interval.
- Condition on total number of shocks to estimate the rate: dependence among trials. Joint distribution not multinomial.

Chi-square test limitations

- Relies on approximation that can be poor.
- Ignores ignores spatial distribution.
- Ignores order of the K intervals: invariant under permutations.
- For instance, the chi-square statistic would have the same value for counts $(N_k) = (3, 1, 0, 2, 0, 4, 1, 0)$ as for counts $(N_k) = (0, 0, 0, 1, 1, 2, 3, 4)$. The latter hardly looks Poisson.
- Hence, chi-square has low power against some alternatives.

Chi-square test limitations

- Relies on approximation that can be poor.
- Ignores ignores spatial distribution.
- Ignores order of the K intervals: invariant under permutations.
- For instance, the chi-square statistic would have the same value for counts $(N_k) = (3, 1, 0, 2, 0, 4, 1, 0)$ as for counts $(N_k) = (0, 0, 0, 1, 1, 2, 3, 4)$. The latter hardly looks Poisson.
- Hence, chi-square has low power against some alternatives.

Chi-square test limitations

- Relies on approximation that can be poor.
- Ignores ignores spatial distribution.
- Ignores order of the K intervals: invariant under permutations.
- For instance, the chi-square statistic would have the same value for counts $(N_k) = (3, 1, 0, 2, 0, 4, 1, 0)$ as for counts $(N_k) = (0, 0, 0, 1, 1, 2, 3, 4)$. The latter hardly looks Poisson.
- Hence, chi-square has low power against some alternatives.

Chi-square test limitations

- Relies on approximation that can be poor.
- Ignores ignores spatial distribution.
- Ignores order of the K intervals: invariant under permutations.
- For instance, the chi-square statistic would have the same value for counts $(N_k) = (3, 1, 0, 2, 0, 4, 1, 0)$ as for counts $(N_k) = (0, 0, 0, 1, 1, 2, 3, 4)$. The latter hardly looks Poisson.
- Hence, chi-square has low power against some alternatives.

KS Test

- Kolmogorov-Smirnov (KS) test better against some alternatives. Test whether, conditional on the number of events, re-scaled times are iid $U[0,1]$.

$$\text{KS statistic } (U[0,1] \text{ null}): D_n = \sup_t \left| \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}(t_i \leq t) - t \right| . \tag{1}$$

- Doesn't require estimating parameters or ad hoc K, B, d, $\hat{\lambda}$.
- Massart-Dvoretzky-Kiefer-Wolfowitz: If null is true,

$$P(D_n > x) \leq 2 \exp\left(-2nx^2\right). \tag{2}$$

Conservative P-values.

KS Test

- Kolmogorov-Smirnov (KS) test better against some alternatives. Test whether, conditional on the number of events, re-scaled times are iid $U[0, 1]$.

$$\text{KS statistic } (U[0,1] \text{ null}): D_n = \sup_t \left| \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}(t_i \leq t) - t \right| . \quad (1)$$

- Doesn't require estimating parameters or ad hoc K, B, d, $\hat{\lambda}$.
- Massart-Dvoretzky-Kiefer-Wolfowitz: If null is true,

$$P(D_n > x) \leq 2 \exp\left(-2nx^2\right). \quad (2)$$

Conservative P-values.

KS Test

- Kolmogorov-Smirnov (KS) test better against some alternatives. Test whether, conditional on the number of events, re-scaled times are iid $U[0,1]$.

$$\text{KS statistic } (U[0,1] \text{ null}): D_n = \sup_t \left| \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}(t_i \leq t) - t \right| . \quad (1)$$

- Doesn't require estimating parameters or ad hoc K, B, d, $\hat{\lambda}$.
- Massart-Dvoretzky-Kiefer-Wolfowitz: If null is true,

$$P(D_n > x) \leq 2 \exp\left(-2nx^2\right). \quad (2)$$

Conservative P-values.

Tests on simulated data

Process	KS power	Chi-square test power
Heterogeneous Poisson	1	0.1658
Gamma renewal	0.0009	1

Estimated power of level-0.05 tests of homogeneous Poisson null hypothesis from 10,000 simulations. Chi-square test uses 10-day intervals, $B = 4$ bins, and $d = B - 2 = 2$ degrees of freedom. “Heterogeneous Poisson”: rate 0.25 per ten days for 20 years, then at rate 0.5 per ten days for 20 years. “Gamma renewal”: inter-event times iid gamma with shape 2 and rate 1.

Methods tested on SCEC data

- **Method 1:** Remove every event in the window of some other event.
- **Method 2:** Divide the catalog into clusters: include an event in a cluster if and only if it occurred within the window of at least one other event in the cluster. In every cluster, remove all events except the largest.
- **Method 3:** Consider the events in chronological order. If the ith event falls within the window of a preceding larger shock that has not already been deleted, delete it. If a larger shock falls within the window of the ith event, delete the ith event. Otherwise, retain the ith event.

Methods 1 and 2 are linked-window methods; Method 3 is main shock method.

Methods tested on SCEC data

- **Method 1:** Remove every event in the window of some other event.
- **Method 2:** Divide the catalog into clusters: include an event in a cluster if and only if it occurred within the window of at least one other event in the cluster. In every cluster, remove all events except the largest.
- **Method 3:** Consider the events in chronological order. If the *i*th event falls within the window of a preceding larger shock that has not already been deleted, delete it. If a larger shock falls within the window of the *i*th event, delete the *i*th event. Otherwise, retain the *i*th event.

Methods 1 and 2 are linked-window methods; Method 3 is main shock method.

Methods tested on SCEC data

- **Method 1:** Remove every event in the window of some other event.
- **Method 2:** Divide the catalog into clusters: include an event in a cluster if and only if it occurred within the window of at least one other event in the cluster. In every cluster, remove all events except the largest.
- **Method 3:** Consider the events in chronological order. If the *i*th event falls within the window of a preceding larger shock that has not already been deleted, delete it. If a larger shock falls within the window of the *i*th event, delete the *i*th event. Otherwise, retain the *i*th event.

Methods 1 and 2 are linked-window methods; Method 3 is main shock method.

SCEC Catalog of Southern California Seismicity $M \geq 3.8$, 1932–1971

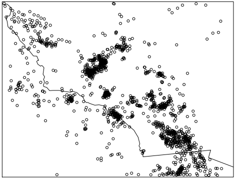


Figure 1: Raw, 1,556 events

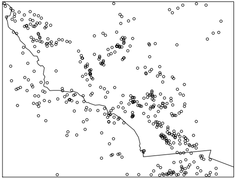


Figure 2: GK linked windows, 424 events

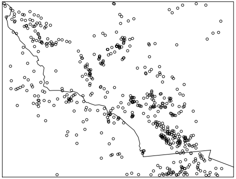


Figure 3: GK main-shock windows, 544 events

Comparison with GK: SCEC Catalog 1932–1971, $M \geq 3.8$

GK used Method 1 but found similar results using Method 2. Tested using using a variety of bin widths. None of their tests rejected. Some difference between their catalog and SCEC.

C.f. KS test plus chi-square test, $B = 4$ and $d = 2$. Reject if either test gave a P-value of less than 0.025; Bonferroni gives level ≤ 0.05.

Method	KS P-value	Chi-square P-value	MLE chi-square P-value	Reject?
1	0.012	0.087	0.087	Yes
2	0.0064	0.297	0.295	Yes
3	0.022	6×10^{-6}	4×10^{-6}	Yes

Distribution of times (after declustering) doesn't seem Poisson.

Exchangeability of times

- Spatially heterogeneous, temporally homogeneous Poisson process (SHTHPP): marginal distribution of times is Poisson, so previous tests reject.
- For SHTHPPs, two events can be arbitrarily close. Window declustering imposes minimum spacing, so can't be SHTHPP.
- For SHTHPPs, conditional on the number of events, the events are iid with probability density proportional to the space-time rate. Conditional on the locations, the marginal distribution of times is iid, hence exchangeable.

Exchangeability of times

- Spatially heterogeneous, temporally homogeneous Poisson process (SHTHPP): marginal distribution of times is Poisson, so previous tests reject.
- For SHTHPPs, two events can be arbitrarily close. Window declustering imposes minimum spacing, so can't be SHTHPP.
- For SHTHPPs, conditional on the number of events, the events are iid with probability density proportional to the space-time rate. Conditional on the locations, the marginal distribution of times is iid, hence exchangeable.

Exchangeability of times

- Spatially heterogeneous, temporally homogeneous Poisson process (SHTHPP): marginal distribution of times is Poisson, so previous tests reject.
- For SHTHPPs, two events can be arbitrarily close. Window declustering imposes minimum spacing, so can't be SHTHPP.
- For SHTHPPs, conditional on the number of events, the events are iid with probability density proportional to the space-time rate. Conditional on the locations, the marginal distribution of times is iid, hence exchangeable.

Exchangeability, contd.

Location of the ith event is (x_i, y_i), $i = 1, \ldots, n$.
x_i is longitude, y_i is latitude.

T_i: Time of the event at (x_i, y_i).

Π: Set of all $n!$ permutations of $\{1, \ldots, n\}$.

Process has *exchangeable times* if, conditional on the locations,

$$\{T_1, \ldots, T_n\} \stackrel{d}{=} \{T_{\pi(1)}, \ldots, T_{\pi(n)}\} \tag{3}$$

for all permutations $\pi \in \Pi$.

Exchangeability, contd.

Location of the ith event is (x_i, y_i), $i = 1, \ldots, n$.
x_i is longitude, y_i is latitude.

T_i: Time of the event at (x_i, y_i).

Π: Set of all $n!$ permutations of $\{1, \ldots, n\}$.

Process has *exchangeable times* if, conditional on the locations,

$$\{T_1, \ldots, T_n\} \stackrel{d}{=} \{T_{\pi(1)}, \ldots, T_{\pi(n)}\} \tag{3}$$

for all permutations $\pi \in \Pi$.

Exchangeability, contd.

Location of the ith event is (x_i, y_i), $i = 1, \ldots, n$.
x_i is longitude, y_i is latitude.

T_i: Time of the event at (x_i, y_i).

Π: Set of all $n!$ permutations of $\{1, \ldots, n\}$.

Process has *exchangeable times* if, conditional on the locations,

$$\{T_1, \ldots, T_n\} \stackrel{d}{=} \{T_{\pi(1)}, \ldots, T_{\pi(n)}\} \quad (3)$$

for all permutations $\pi \in \Pi$.

Exchangeability, contd.

Location of the ith event is (x_i, y_i), $i = 1, \ldots, n$.
x_i is longitude, y_i is latitude.

T_i: Time of the event at (x_i, y_i).

Π: Set of all $n!$ permutations of $\{1, \ldots, n\}$.

Process has *exchangeable times* if, conditional on the locations,

$$\{T_1, \ldots, T_n\} \stackrel{d}{=} \{T_{\pi(1)}, \ldots, T_{\pi(n)}\} \tag{3}$$

for all permutations $\pi \in \Pi$.

Exchangeability, contd.

- SHTHPP has exchangeable times.
- If events close in space tend to be close in time—the kind of clustering real seismicity exhibits—times not exchangeable.
- If events close in space tend to be distant in time—e.g., from window methods for declustering—times not exchangeable.

Exchangeability, contd.

- SHTHPP has exchangeable times.
- If events close in space tend to be close in time—the kind of clustering real seismicity exhibits—times not exchangeable.
- If events close in space tend to be distant in time—e.g., from window methods for declustering—times not exchangeable.

Exchangeability, contd.

- SHTHPP has exchangeable times.
- If events close in space tend to be close in time—the kind of clustering real seismicity exhibits—times not exchangeable.
- If events close in space tend to be distant in time—e.g., from window methods for declustering—times not exchangeable.

Permutation test set up

- $\hat{P}_n$: empirical distribution of the times and locations of the n observed events.
- $\tau(\hat{P}_n)$: projection of $\hat{P}_n$ onto the set of distributions with exchangeable times
  τ puts equal mass at every element of the orbit of data under the permutation group on times.
- $V \subset R^3$ is a *lower-left quadrant* if:

$$V\{x = (x, y, t) \in R^3 : x \leq x_0 \text{ and } y \leq y_0 \text{ and } t \leq t_0\}. \quad (4)$$

- **V**: the set of all lower-left quadrants.

Permutation test set up

- $\hat{P}_n$: empirical distribution of the times and locations of the n observed events.
- $\tau(\hat{P}_n)$: projection of $\hat{P}_n$ onto the set of distributions with exchangeable times
  τ puts equal mass at every element of the orbit of data under the permutation group on times.
- $V \subset R^3$ is a *lower-left quadrant* if:

$$V\{x = (x, y, t) \in R^3 : x \le x_0 \text{ and } y \le y_0 \text{ and } t \le t_0\}. \quad (4)$$

- **V**: the set of all lower-left quadrants.

Permutation test set up

- $\hat{P}_n$: empirical distribution of the times and locations of the n observed events.
- $\tau(\hat{P}_n)$: projection of $\hat{P}_n$ onto the set of distributions with exchangeable times
  τ puts equal mass at every element of the orbit of data under the permutation group on times.
- $V \subset R^3$ is a *lower-left quadrant* if:

$$V\{x = (x, y, t) \in R^3 : x \leq x_0 \text{ and } y \leq y_0 \text{ and } t \leq t_0\}. \quad (4)$$

- **V**: the set of all lower-left quadrants.

Permutation test set up

- $\hat{P}_n$: empirical distribution of the times and locations of the n observed events.
- $\tau(\hat{P}_n)$: projection of $\hat{P}_n$ onto the set of distributions with exchangeable times
  τ puts equal mass at every element of the orbit of data under the permutation group on times.
- $V \subset R^3$ is a *lower-left quadrant* if:

$$V\{x = (x, y, t) \in R^3 : x \leq x_0 \text{ and } y \leq y_0 \text{ and } t \leq t_0\}. \quad (4)$$

- **V**: the set of all lower-left quadrants.

Test statistic

$$\sup_{V \in \mathbf{V}} |\hat{P}_n(V) - \tau(\hat{P}_n)(V)| \tag{5}$$

- Generalization of the KS statistic to three dimensions.
- Suffices to search a finite subset of $\mathbf{V}$.
  Can sample at random from that finite subset for efficiency.
- Calibrate by simulating from $\tau(\hat{P}_n)$—permuting the times (Romano)

Test statistic

$$\sup_{V \in \mathbf{V}} |\hat{P}_n(V) - \tau(\hat{P}_n)(V)| \tag{5}$$

- Generalization of the KS statistic to three dimensions.
- Suffices to search a finite subset of **V**.
  Can sample at random from that finite subset for efficiency.
- Calibrate by simulating from $\tau(\hat{P}_n)$—permuting the times (Romano)

Test statistic

$$\sup_{V \in \mathbf{V}} |\hat{P}_n(V) - \tau(\hat{P}_n)(V)| \quad (5)$$

- Generalization of the KS statistic to three dimensions.
- Suffices to search a finite subset of **V**.
  Can sample at random from that finite subset for efficiency.
- Calibrate by simulating from $\tau(\hat{P}_n)$—permuting the times (Romano)

Results of exchangeability test: 2009 SCEC data $M \geq 2.5$

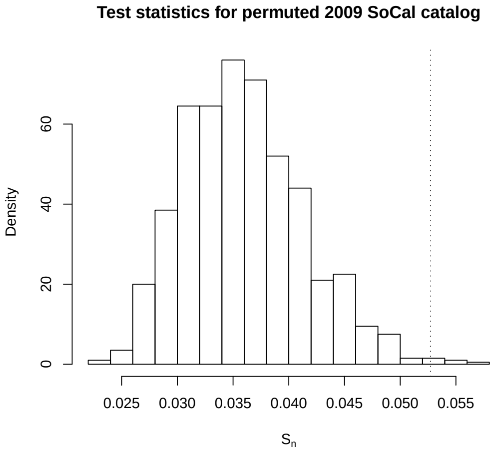


Figure 4: Reasenberg declustering: 475 events. 1-tailed P-value ≈ 0.003 (99% CI $[0.0003, 0.011]$). For raw catalog of 753 events, test statistic is larger than any of the statistics for 1,000 permuted catalogs: $P < 0.001$ (99% CI $[0, 0.007]$).

Discussion: Seismology

- Declustered catalogs don't look Poisson in time.
- Declustered catalogs can't be Poisson in space-time.
- Declustered catalogs don't seem to have exchangeable times, necessary condition for Poisson.
- No clear definition of foreshock, main shock, aftershock.
- All big shocks can cause damage and death. Physics doesn't distinguish main shocks from others. So why decluster?

Discussion: Seismology

- Declustered catalogs don't look Poisson in time.
- Declustered catalogs can't be Poisson in space-time.
- Declustered catalogs don't seem to have exchangeable times, necessary condition for Poisson.
- No clear definition of foreshock, main shock, aftershock.
- All big shocks can cause damage and death. Physics doesn't distinguish main shocks from others. So why decluster?

Discussion: Seismology

- Declustered catalogs don't look Poisson in time.
- Declustered catalogs can't be Poisson in space-time.
- Declustered catalogs don't seem to have exchangeable times, necessary condition for Poisson.
- No clear definition of foreshock, main shock, aftershock.
- All big shocks can cause damage and death. Physics doesn't distinguish main shocks from others. So why decluster?

Discussion: Seismology

- Declustered catalogs don't look Poisson in time.
- Declustered catalogs can't be Poisson in space-time.
- Declustered catalogs don't seem to have exchangeable times, necessary condition for Poisson.
- No clear definition of foreshock, main shock, aftershock.
- All big shocks can cause damage and death. Physics doesn't distinguish main shocks from others. So why decluster?

Discussion: Seismology

- Declustered catalogs don't look Poisson in time.
- Declustered catalogs can't be Poisson in space-time.
- Declustered catalogs don't seem to have exchangeable times, necessary condition for Poisson.
- No clear definition of foreshock, main shock, aftershock.
- All big shocks can cause damage and death. Physics doesn't distinguish main shocks from others. So why decluster?

Discussion: Statistics

- The test matters. What's the scientific question?
- Trivial to make declustering method pass test: Delete most of the events.
- Suggests optimization problem: remove fewest events to pass test.
- Combinatorially complex in general.
- For test statistic like $\sup_{V \in \mathbf{V}} |\hat{P}_n(V) - \tau(\hat{P}_n)(V)|$, relaxation is linear program.
- Interesting? No obvious practical import.

Discussion: Statistics

- The test matters. What's the scientific question?
- Trivial to make declustering method pass test: Delete most of the events.
- Suggests optimization problem: remove fewest events to pass test.
- Combinatorially complex in general.
- For test statistic like $\sup_{V \in \mathbf{V}} |\hat{P}_n(V) - \tau(\hat{P}_n)(V)|$, relaxation is linear program.
- Interesting? No obvious practical import.

Discussion: Statistics

- The test matters. What's the scientific question?
- Trivial to make declustering method pass test: Delete most of the events.
- Suggests optimization problem: remove fewest events to pass test.
- Combinatorially complex in general.
- For test statistic like $\sup_{V \in \mathbf{V}} |\hat{P}_n(V) - \tau(\hat{P}_n)(V)|$, relaxation is linear program.
- Interesting? No obvious practical import.

Discussion: Statistics

- The test matters. What's the scientific question?
- Trivial to make declustering method pass test: Delete most of the events.
- Suggests optimization problem: remove fewest events to pass test.
- Combinatorially complex in general.
- For test statistic like $\sup_{V \in \mathbf{V}} |\hat{P}_n(V) - \tau(\hat{P}_n)(V)|$, relaxation is linear program.
- Interesting? No obvious practical import.

Discussion: Statistics

- The test matters. What's the scientific question?
- Trivial to make declustering method pass test: Delete most of the events.
- Suggests optimization problem: remove fewest events to pass test.
- Combinatorially complex in general.
- For test statistic like $\sup_{V \in \mathbf{V}} |\hat{P}_n(V) - \tau(\hat{P}_n)(V)|$, relaxation is linear program.
- Interesting? No obvious practical import.

Discussion: Statistics

- The test matters. What's the scientific question?
- Trivial to make declustering method pass test: Delete most of the events.
- Suggests optimization problem: remove fewest events to pass test.
- Combinatorially complex in general.
- For test statistic like $\sup_{V \in \mathbf{V}} |\hat{P}_n(V) - \tau(\hat{P}_n)(V)|$, relaxation is linear program.
- Interesting? No obvious practical import.